ECONSTOR Make Your Publications Visible.

A Service of

ZBW

Leibniz-Informationszentrum Wirtschaft Leibniz Information Centre for Economics

Bensch, Gunther; Rose, Julian; Neubauer, Florian; Ankel-Peters, Jörg; Brodeur, Abel

Working Paper

Report on "Partisan effects of information campaigns in competitive authoritarian elections: evidence from Bangladesh" by Ahmed et al. (2024)

I4R Discussion Paper Series, No. 216

Provided in Cooperation with:

The Institute for Replication (I4R)

Suggested Citation: Bensch, Gunther; Rose, Julian; Neubauer, Florian; Ankel-Peters, Jörg; Brodeur, Abel (2025) : Report on "Partisan effects of information campaigns in competitive authoritarian elections: evidence from Bangladesh" by Ahmed et al. (2024), I4R Discussion Paper Series, No. 216, Institute for Replication (I4R), s.l.

This Version is available at: https://hdl.handle.net/10419/314328

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



WWW.ECONSTOR.EU

INSTITUTE for

No. 216 DISCUSSION PAPER SERIES

Report on "Partisan effects of information campaigns in competitive authoritarian elections: evidence from Bangladesh" by Ahmed et al. (2024)

Gunther Bensch Florian Neubauer Abel Brodeur Julian Rose Jörg Ankel-Peters

This paper received responses:

Ahmed, F., R. Hodler, and A. Islam. 2025. Responses to Bensch, Rose, Neubauer, Ankel-Peters, and Brodeur. *IAR Discussion Paper Series* No. 217. Institute for Replication

March 2025



I4R DISCUSSION PAPER SERIES

I4R DP No. 216

Report on "Partisan effects of information campaigns in competitive authoritarian elections: evidence from Bangladesh" by Ahmed et al. (2024)

Gunther Bensch¹, Julian Rose^{1,2}, Florian Neubauer¹, Jörg Ankel-Peters^{1,3}, Abel Brodeur^{4,3}

¹RWI – Leibniz Institute for Economic Research, Essen/Germany ²Ludwig-Maximilians-University Munich/Germany ³Institute for Replication ⁴University of Ottawa/Canada

MARCH 2025

Any opinions in this paper are those of the author(s) and not those of the Institute for Replication (I4R). Research published in this series may include views on policy, but I4R takes no institutional policy positions.

I4R Discussion Papers are research papers of the Institute for Replication which are widely circulated to promote replications and metascientific work in the social sciences. Provided in cooperation with EconStor, a service of the ZBW - Leibniz Information Centre for Economics, and RWI - Leibniz Institute for Economic Research, I4R Discussion Papers are among others listed in RePEc (see IDEAS, EconPapers). Complete list of all I4R DPs - downloadable for free at the I4R website.

I4R Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

Editors							
Abel Brodeur University of Ottawa	Anna Dreber Stockholm Sch	ool of Economics	Jörg Ankel-Peters RWI – Leibniz Institute for Economic Research				
E-Mail: joerg.peters@rwi-es RWI – Leibniz Institute for E	<u>ssen.de</u> Economic Research	Hohenzollernstraß 45128 Essen/Gerr	e 1-3 <u>www.i4replication.o</u>	rg			

ISSN: 2752-1931

Report on "Partisan effects of information campaigns in competitive authoritarian elections: evidence from Bangladesh" by Ahmed et al. (2024)

Gunther Bensch, Julian Rose, Florian Neubauer, Jörg Ankel-Peters & Abel Brodeur*

March 18, 2025

*Bensch: RWI – Leibniz Institute for Economic Research, Rose: RWI – Leibniz Institute for Economic Research & Ludwig-Maximilians-University Munich, Neubauer: RWI – Leibniz Institute for Economic Research, Ankel-Peters: RWI – Leibniz Institute for Economic Research & Institute for Replication, Brodeur: University of Ottawa & Institute for Replication (abrodeur@uottawa.ca).

Abstract

This report compiles our recent comment on Ahmed, Hodler, and Islam (2024, AHI-2024) and our response to the authors' reply to our comment. Our report is one element in a concerted forensic reproduction of studies based on data collected by GDRI, a Bangladesh-based survey company. We appreciate the authors' acknowledgment of discrepancies in AHI-2024. These are consequential admissions given the forensic nature of this report. We also clarify that these discrepancies are likely to drive AHI-2024's main results. Furthermore, our response shows that the authors' reply contains new contradictions. Overall, our investigation of the paper, the study's documentation, and its replication package raise serious concerns about the integrity of the data and the study design.

To facilitate access to the debate, this report features our <u>response</u> to the original authors' reply first (<u>the authors' reply can be found here</u>), followed by our initial <u>comment</u> on AHI-2024. The original authors also replied to our response (<u>which can be found here</u>).

A response to Ahmed et al. (2025)

GUNTHER BENSCH, JULIAN ROSE, FLORIAN NEUBAUER, JÖRG ANKEL-PETERS & ABEL BRODEUR*

March 07, 2025

*Bensch: RWI – Leibniz Institute for Economic Research, Rose: RWI – Leibniz Institute for Economic Research & Ludwig-Maximilians-University Munich, Neubauer: RWI – Leibniz Institute for Economic Research, Ankel-Peters: RWI – Leibniz Institute for Economic Research & Institute for Replication, Brodeur: University of Ottawa & Institute for Replication (abrodeur@uottawa.ca).

1. Summary

We thank Ahmed, Hodler, and Islam (2025, henceforth AHI-2025) for their polite and professional response to our comment (Bensch et al. 2025, henceforth BRNAPB-2025). We also appreciate their acknowledgment of mistakes in Ahmed, Hodler, and Islam (2024, henceforth AHI-2024). What is still pending is the corrected version of the erroneous map, which AHI-2025 announce to "provide [..] in due course". In this note, we nevertheless briefly comment on the preliminary response and contextualize the purpose of BRNAPB-2025 as one element in a concerted forensic reproduction of studies based on data collected by GDRI, a Bangladesh-based survey company.

To recapitulate, in BRNAPB-2025 we document one peculiarity in the AHI-2024 data as well as two major discrepancies in AHI-2024's description of how the data was collected and a flawed map depicting the sample of villages and their treatment and classification status. AHI-2025 contest the peculiarity, but they concede the discrepancies and the flawed map. As we will argue in the next section, these are consequential admissions given the nature of this reproduction. AHI-2025's explanation of these admissions contains new contradictions. Moreover, we further clarify our view on the *extremely* impressive data collection effort underlying the AHI-2024 study (Section 3) and briefly contextualize AHI-2025's admissions of the discrepancies (Section 4). Not least, AHI-2025's discussion of the peculiarity described in BRNAPB-2025 absorbs most of the authors' response, but without linking it to the admitted discrepancies. We therefore briefly reiterate this link in Section 5. Overall, our investigation of the paper, the study's documentation, and its replication package raises serious concerns about the integrity of the data and the study design.

Institute for Replication

I4R DP No. 216

2. A forensic reproduction

The aim of BRNAPB-2025 is not to discuss the interpretation of AHI-2024. Rather, BRNAPB-2025 is part of a concerted project investigating concerns about data collected and Randomized Controlled Trials implemented by GDRI. Companion comments identified discrepancies in other papers based on GDRI's data related to different stages of the research process including the study design and sampling (Brodeur et al. 2025; Fiala et al. 2025; Kjelsrud et al. 2025). In the meantime, these companion comments have led to strong reactions from some of the authors, including public dissociations and requests to journals to withdraw their names from already published papers. In this context, BRNAPB-2025 is a forensic reproduction that scrutinizes potential discrepancies in the data underlying AHI-2024. It is not a classical robustness reproduction testing the sensitivity of results and raising questions about AHI-2024's interpretations (although such questions might emerge as part of this examination, see Section 5). It is hence not BRNAPB-2025's ambition to judge whether documented issues are consequential for the results. Forensic reproductions intend to provide evidence that significant inconsistencies exist warranting closer scrutiny. In this regard, we note that AHI-2025 highlight a division of labour, very saliently on their first page, with Ahmed and Islam taking responsibility "for the RCT and the data collection in Bangladesh" and Hodler taking responsibility only for having taken "the lead in analyzing the data and writing the paper".

3. AHI-2024's data collection

In BRNAPB-2025, we refer to the data collection underlying AHI-2024 as "impressive" (p.3), which AHI-2025 quote in their response. We would therefore like to clarify that the footnote in our comment on the data collection is supposed to convey that especially the elicitation of the main outcome variable, voting, appears to us highly implausible. To obtain an objective measure for 'voting', the authors¹ visited all households in their panel on the evening of election day or the day after, in order to visually inspect inked fingers alongside a short interview. Hence, within a few hours and under extreme time pressure 11,843 households were visited and 22,686 respondents interviewed (two spouses in each household). This is

¹ AHI-2024 explicitly describe the engagement with GDRI and state that most of the field activities were implemented by the company. Beyond this general description it is not distinguishable though, which activity was decidedly implemented by GDRI and which by the authors. Throughout the paper, the authors say "we", for example "we delivered the treatment messages in three different ways" (p.1314).

indeed difficult to believe. It is furthermore remarkable that the household attrition rate is at 1% only, that is, virtually all households that had been interviewed prior to the election could also be retrieved within these few hours. And what is even more implausible is that for the retrieved households both spouses were found – in all cases. This is intuitively very surprising, but also not in line with comparable evidence. A similar study in Pakistan, also on postelection inspection of inked fingers, reports considerable rates of untraceable respondents (Giné and Mansuri 2018). Likewise, studies on other topics but involving simultaneous spouse interviewing encounter similarly high non-response rates for one spouse, even with much less time pressure than in the present case (see e.g. Kilic and Moylan 2016; Masselus and Fiala 2024). Note that AHI-2024 do not mention any imputation of missing data.

4. The uncontested discrepancies

In BRNAPB-2025, we document two major discrepancies in AHI-2024's description of how the data was collected and, in addition, we reveal that AHI-2024's map depicting the sample of villages and their treatment and classification status is flawed. We could identify these two documented discrepancies because the Pre-Analysis Plan (PAP), the discussion paper version published in the renowned CEPR series, and AHI-2024, the *Economic Journal* version, are contradictory. These contradictions are substantive, since they relate to crucial elements of the data collection process – sampling and village classification – that are decisive for AHI-2024's main findings (see next section).

AHI-2025 frame these discrepancies as "refinements" (p.2) that occurred during the implementation of the study and thereby turn the spotlight on deviations from the PAP, arguing that it "is not uncommon in such cases to refine the initial strategy in the field" (p.2). We agree with this general statement, and it is common sense in the profession on how to handle PAPs in practice (see for example Banerjee et al. 2020).² However, as the AEA registry documents, these "refinements" were not made post-PAP and pre-implementation of the study. The PAP was registered on December 14, 2018 – when the data collection had already been ongoing and hence both sampling and village classification decisions had already been

² It is concerning that such substantive deviations from the PAP are not explicitly mentioned in AHI-2025, but this is not part of our argument.

made.³ During this registration on December 14, the authors posted the sampling procedure into the registry template that is at odds with the description in AHI-2024 ("We then randomly select 35-40 individuals per village, with roughly equal gender balance"). Additional changes to the registry/PAP were made in January 2024, but only the final number of villages were updated.

The descriptions of the sampling and village classification procedures were only changed prior to or during the submission process, as it is evidenced by the CEPR discussion paper featuring the same description as the PAP. AHI-2025 concede this discrepancy between AHI-2024 and the CEPR discussion paper version. We find two of the three explanations provided in the response unconvincing. First, the CEPR version contained "condensed" (p.2) information as it was foreseen for a short-paper submission and, second, "descriptions [...] became longer in response to specific comments raised during the referee process at the *Economic Journal*" (p.2). However, the descriptions in all versions are as short as they are substantive. The longer version in the *Economic Journal* consists of 23 words (on sampling) and 14 words (on classification), respectively, the "condensed" CEPR version consists of 8 (sampling) and 7 words (classification), respectively.⁴ Also note that the text in the CEPR version slightly differs from the PAP, indicating that it was not merely copy and pasted. The third explanation provided by AHI-2025 is more plausible and refers to the "division of labor (see Section 1)" (p.2) between Hodler on the one hand and Ahmad and Islam on the other (see above) and to a "less-than-perfect information flow within the team" (p.2).

Of course, information flow can be less than perfect, and this would help to understand how a wrong description of *minor* importance proliferates through the entire study documentation. However, these descriptions are of <u>utmost</u> importance for AHI-2024's results, as AHI-2025 explain themselves (see again next section).

³ According to AHI-2024, the first survey was implemented "about six weeks prior to the election", and hence in mid-November 2018 and "The second pre-treatment survey was conducted from December 12-16" (p.1315).

⁴ Core statements on sampling are "Third, we randomly selected 40 households per village, [...]" in the CEPR version: versus "[...] we selected a neighbourhood (para) in each village. This selection was non-random as we intended to target individuals in government and opposition strongholds." in AHI-2024. And on classification "For polling stations consisting of multiple villages [...]" in CEPR version versus "If the relative electoral strength of the two parties changed from 2001 to 2008 [...]" in AHI-2024.

The contradictions between the CEPR discussion paper and AHI-2024 are also related to the map displaying the entire sample of villages, their treatment status and classification – prominently depicted as Figure 1 in AHI-2024. This map, as AHI-2025 concede, "contains errors" (p.7). They further note that they "are currently investigating the issue and will provide a revised map in due course", referring to the "limited time available" due to the timeframe set by the I4R. We are confident that generating the corrected map only requires a manageable amount of time and are looking forward to receiving it soon to assess whether these mapping inaccuracies indeed do not affect their results as claimed in AHI-2025.

5. The peculiarity reiterated

AHI-2025 dedicate most of their response to the discussion of the peculiarity outlined in BRNAPB-2025, and they name their sampling procedure as the major reason explaining the peculiarity – without making the link to the discrepancies discussed in the previous section. We therefore reiterate that our Figure 1 in BRNAPB-2025 demonstrates the importance of the sampling procedure for the treatment effects in AHI-2024 that the authors contradictorily describe in PAP, the CEPR version and AHI-2024 (see previous section). Figure 1 is hence important to understand that these contradictions are not about "refinements" but consequential changes in the study design that are likely to drive the results.

Beyond this, we merely note that by these additional analyses and discussions, the authors ascribe an immense importance to their within-village sampling strategy. This is, we believe, not appropriately accounted for in the interpretations and inferences made in AHI-2024. The very specific sample of high-partisan neighbourhoods has profound implications for the external validity of their findings – since everything the authors claim only holds for such high-partisan populations. Unfortunately, though, this external validity debate is not the point of our comment and this debate.

6. Conclusion

Throughout the response (AHI-2025) the authors repeatedly apologize for "inaccuracies" and "mistakes" (in the map) as well as the "incomplete and potentially misleading descriptions of our village classification and sampling strategies in our working paper" (p.2) – which is genuinely appreciated. But AHI-2025 also repeatedly emphasize that they do not affect

"results or their interpretation" (e.g., p. 7). This is a misconception of the purpose of BRNAPB-2025. We have now clarified that our reproduction is of forensic nature and the discrepancies in how crucial study features are presented throughout the study documentation need to be interpreted within the context of the wider concerns surrounding the quality of data and randomizations implemented by GDRI.

A comment on "Partisan effects of information campaigns in competitive authoritarian elections: evidence from Bangladesh" by Ahmed et al. (2024)

GUNTHER BENSCH, JULIAN ROSE, FLORIAN NEUBAUER, JÖRG ANKEL-PETERS & ABEL BRODEUR*

February 2025

*Bensch: RWI – Leibniz Institute for Economic Research, Rose: RWI – Leibniz Institute for Economic Research & Ludwig-Maximilians-University Munich, Neubauer: RWI – Leibniz Institute for Economic Research, Ankel-Peters: RWI – Leibniz Institute for Economic Research & Institute for Replication, Brodeur: University of Ottawa & Institute for Replication (abrodeur@uottawa.ca).

Abstract

Ahmed, Hodler, and Islam (2024, henceforth AHI) conduct a village-level Randomized Controlled Trial and examine the impact of get-out-the-vote (GOTV) campaigns on voter turnout during the 2018 Bangladesh election. AHI's main finding is a heterogeneity in the responsiveness to the treatment. This heterogeneity analysis relies on a classification of villages into either government-leaning or opposition-leaning villages. The classification is based on historical voting behaviour or focus group discussions conducted in the villages. We successfully reproduce all results in AHI but find surprising patterns with respect to the main finding. We focus on a variable based on the AHI data that is conceptually close to AHI's classification criterion, partisanship strength, and find that it affects the control and treatment groups in fundamentally different ways. Partisanship strength mediates the treatment effects at a sharp cut-off. These patterns remain unexplored in the original study. Moreover, we document inconsistencies in the description of two pivotal aspects of the paper – the sampling strategy and village classification – between the pre-analysis plan, the discussion paper version and AHI, the journal version. We also show that the map displaying the sampled villages in AHI does not align with the data in the reproduction package. These peculiarities warrant further clarification.

1. Introduction

Ahmed, Hodler, and Islam (2024, henceforth AHI) conduct a large-scale Randomized Controlled Trial on the village level to examine the impact of non-partisan get-out-the-vote (GOTV) campaigns on voter turnout during the 2018 Bangladeshi general election. AHI employ two treatments with different GOTV framings, a *policy treatment* emphasizing towards treated participants that higher voter turnout influences later policy decisions, and a *legitimacy treatment* emphasizing that higher voter turnout enhances the legitimacy of the winning party's government. Voter turnout in the 2018 election (henceforth: voting) is the main outcome. Our paper is a computational and robustness reproduction of AHI. We successfully reproduce all results in AHI. Scope for a robustness reproduction is limited because of a comprehensive yet parsimonious replication package that only contains the analysis data.

AHI's main finding is a heterogeneity in the responsiveness to the treatment according to whether villages lean towards the currently ruling party, i.e. the government, or the opposition. AHI's abstract states: *"Both treatments increase turnout (measured by ink marks) in government strongholds, but decrease turnout in opposition strongholds"* (AHI, p.1303). This result is based on a village classification by AHI. During our reproduction some peculiarities related to this classification emerged, which are presented in this report.

AHI classified the villages into government- or opposition-leaning villages prior to data collection, for most villages based on which party received more votes in both the 2001 and 2008 elections. This criterion applied to two thirds of villages. For the remaining third of villages, the criterion is inconclusive, according to AHI because the majority vote switches between 2001 and 2008. This remaining third, AHI classified based on focus group discussions conducted by their field staff. AHI then use this classification to stratify the sample into two equally sized groups of villages. This stratification underlies the heterogeneity prominently featured in AHI's abstract: there is no overall treatment effect but voting increased by 7–15 percentage points in government-leaning villages and decreased by 10–21 percentage points in opposition-leaning villages.

Following the village classification and the stratified randomization, AHI collected four rounds of household-level data: two rounds prior to the 2018 election, one on election day and another one week later. They cover a sample of up to 40 households per village in 302 villages

Institute for Replication

I4R DP No. 216

in total. The randomized intervention was implemented after the second round of data collection. Within villages, AHI selected neighbourhoods with particularly strong support for the dominant party. The key variable, voting, is the result of AHI's impressive data collection effort conducted on or shortly after election day, including the visual verification of ink-marked fingers for nearly 24,000 survey participants.

In this report, prepared for the Institute for Replication, we reveal some peculiar patterns that are not explored by AHI. First, we conduct a graphical analysis of the main outcome variable, voting, at the village level across the village-level vote gap distribution, which reflects the difference in vote shares between the two parties in the 2001 and 2008 elections. The *vote gap* variable enables us to examine treatment responsiveness according to the partisanship strength in each village. We observe intuitive behaviour in the control group, where voting smoothly varies with partisanship strength (see Figure 1). This implies that villages around the zero point for partisanship strength (i.e. those slightly leaning toward either side) behave similarly. The treatment group, in contrast, displays a sharp discontinuity at the zero point, suggesting an abrupt increase of approximately 25 and 40 percentage points in voting for the two treatments when moving from a slightly opposition-leaning village to a slightly government-leaning one. This pattern implies that village classification functions as an extremely sharp threshold. What is more, the treatment response below and above this cut-off seems to be uniform, with no difference between villages that slightly or strongly favour one side. Hence, partisanship strength seems to affect the control and treatment groups in fundamentally different ways - which is possible, but not further discussed in AHI.

The second peculiar pattern we observe is related to the classification underlying the stratification, and hence AHI's main finding. We investigate the classification in two consecutive steps. First, we examine the pre-analysis plan, using the entries in the AEA Registry (Hodler and Islam 2018, 2024; henceforth PAP⁵), and the discussion-paper version (Ahmed et al. 2020; henceforth DP), comparing these to AHI, i.e. the version published in *The Economic Journal*. We identify inconsistencies in the description of the sampling and the village classification between PAP/DP and AHI. Yet, while the textual descriptions for the village classification and treatment

⁵ Our analysis of the PAP is based on the publicly available information in the AEA Registry, with latest update in January 2024. We did not request access to the full PAP document.

status (Fig. 1 in DP and AHI) are identical, as are the reported results. Moreover, we find indication in the data that the map is inaccurate. We confirm these inconsistencies by a deeper analysis of the village classification using additional GPS data sources from the national election data. Our findings may be explained by a reclassification of villages between PAP/DP and AHI. Yet, if this reclassification occurred after the treatment had been delivered, it would not only contradict the descriptions of the study design in both versions of the paper but also conflict with the reported results being identical in the DP and in AHI, and with the presentation in the map.

2. Computational Reproducibility

The reproduction package is publicly available (Ahmed et al. 2023). It contains the analysis files, that is, three do-files and several cleaned datasets used to generate the variables, reproduce the main figures and tables, and create the appendix. It does not include any raw data or cleaning code. We perfectly computationally reproduce the original results by simply inserting the file path and then run the three provided do-files subsequently (see Table 1). The documentation is largely comprehensive, featuring a read-me and additional information on the experimental design, consent form, IRB approval, and treatment delivery. It does not provide any survey instruments.

	Fully	Partial	No
Raw data provided			х
Cleaning code provided			х
Analysis data provided	x		
Analysis code provided	x		
Reproducible from raw data			х
Reproducible from analysis data	x		

Table 1: Result of computational reproduction	
---	--

3. The vote gap cut-off peculiarity

3.1 The vote gap as a measure of partisanship strength

Our investigation focuses on how AHI's primary outcome, voting in the 2018 election, behaves relative to the historical village-level vote gap (hereafter: vote gap). We use the vote gap as a good measure of partisanship strength, which is at the bottom of AHI main finding: the treatment heterogeneity according to whether the villages lean towards the opposition or the government. In Figure 1, the y-axis represents the voting variable⁶ and the x-axis denotes the vote gap – distinguished between the control group and the two treatment groups, policy and legitimacy.

The vote gap reflects the average difference in vote shares at the village level between the government and main opposition party, averaged over the 2001 and 2008 elections. Villages with a negative vote gap (left of the dashed line indicating the zero point for the vote gap) voted more for the opposition party in 2001 and 2008, while those with a positive vote gap (right of the dashed line) favoured the government party. Villages with vote gaps near zero are more likely to have been classified based on focus groups (light blue circles) rather than only election data (dark blue circles), which is plausible given that, according to AHI, villages classified based on focus groups are those villages where voting preference shifted between 2001 and 2008.

3.2 The peculiarity

Figure 1 shows that in the control group, voting gradually increases from strongly oppositionleaning villages to strongly government-leaning villages. Hence, in villages near a vote gap of zero, where partisanship strength is ambiguous, voting is similar across government and opposition villages. In line with expectation, villages classified by focus group discussion are clustered around the vote gap of zero in the control group. These villages reveal somewhat noisier voting behaviour in the 2018 election. This pattern is plausible, as many of these villages have shifted voting in the past elections between government and opposition, making households less likely to act uniformly. In contrast, the treatment groups display a distinctly different pattern. For both treatments, voting remains constant within each village type, with an abrupt discontinuity at the vote gap of zero. This pattern is surprising for two reasons. First,

⁶ According to AHI, voting was measured by asking a direct question. But since AHI deemed voting behavior to be a "sensitive topic [...] in a competitive authoritarian election" (p.1304), in a "logistically challenging data-gathering effort" (p.1304) their enumerators assessed whether respondents have ink-marked fingers. This was elicited for 23,686 respondents (main respondent and their spouses in 11,843 households) in AHI's third round of data collection. AHI conducted this survey round "in the late afternoon and evening of election day to be extra cautious [about the durability of the ink mark]. If we could not complete the work or locate some respondents (or their spouses) at home on election day, we continued the survey on the following day." (p.1315) This impressive effort resulted in an attrition rate of just 1%. In the remaining 11,843 households, ink marks were successfully recorded for both spouses, as there are no missing values for these variables in the reproduction data. All our results in this report are for the main respondent. Results look very similar for spouses and are therefore not presented.

partisanship strength would probably be expected to continuously mediate the strength of the effects in the treatment groups (as it is the case in the control group).





Note: The vote gap is calculated as the difference in vote share between government and opposition party, averaged over two previous elections (2001 and 2008). The solid line trends give the predicted values from a regression of the outcome variable on a second-degree polynomial in the vote gap that uses a triangular kernel and a bandwidth of 50%. For all groups, separate predictions are run by the village classification.

Second, the abrupt discontinuity suggests that households in marginally government-leaning villages behave fundamentally differently from those in marginally opposition-leaning ones. Taking the example of the legitimacy treatment, Figure 1 suggests a sharp increase of approximately 40 percentage points in voting when moving from a slightly opposition-leaning village to a slightly government-leaning one. In line with expectation, most of the villages that were classified based on focus-group discussions can be found around the vote gap of zero. Especially here, some degree of imprecision in terms of classification would be expected. Yet, the pattern in Figure 1 implies that village classification functions as an extremely sharp threshold. All villages respond to the treatments exactly in line with their classification, that is, all villages classified as government-leaning villages show the increase in voting, and all villages classified as opposition-leaning villages show the decrease in voting.

The peculiar pattern observed in Figure 1 is not mentioned in AHI and therefore not explained. One potential factor that may explain part of this peculiarity is AHI's household sampling strategy, which we explore in the next sub-section. The classification of villages, crucial for AHI's main claim, might also play a role. We examine this further in Section 4.

3.3. Household sampling

AHI's household sampling "selected a neighbourhood (para) in each village. This selection was non-random as we intended to target individuals in government and opposition strongholds" (AHI, p.1311). The description suggests that within government (opposition) villages, AHI selected neighbourhoods that were particularly strong leaning towards the government (opposition). This might explain the peculiar results in the treatment group presented in the previous section. However, this description contrasts with the sampling description in the PAP, an updated PAP after completion of data collection⁷, and the DP, where the sampling is explicitly described as simple random sampling at village level (see Table 2 for the relevant excerpts from each source).

Pre-analysis plan (PAP)	Discussion paper (DP)	Published paper (AHI)			
We randomly select 35-40 individuals per village, with	Third, we randomly selected 40 households per village, with a	Second, we selected a neighbourhood (para) in each village. This selection was			
roughly equal gender	focus on the married and literate	non-random as we intended to target			
balance. These are the	individuals between the age of	individuals in government and opposition			
individuals we survey twice	20 and 55 years (who can read	strongholds. We used our village- level			
prior to the general elections	and understand our treatment	focus-group discussions to learn about the			
and thereafter. In addition,	messages). We surveyed these	neighbourhoods of government			
we deliver the respective	households with gender balance	(opposition) villages with a particularly			
treatment messages to these	among respondents, who are	strong leaning towards the AL (BNP). Third,			
35-40 individuals in the	also our target subjects for the	we relied on systematic random sampling			
treatment villages.	treatments. (DP, p.7)	to select 40 individuals from within each of these neighbourhoods. (AHI, p. 1311)			

The sampling strategy might have been changed after the PAP and before going to the field, but this explanation is not consistent with the description in the DP. The potentially revised sampling strategy would also lead to a discontinuity in the control group – which is perceivable as a nuance but not very pronounced. More importantly, if this change in the sampling strategy indeed explained the patterns in Figure 1, it would in our view significantly alter the theoretical claim and the key takeaways of AHI. It is therefore surprising that this

⁷ Note that AHI updated the pre-registration on the AEA RCT Registry website on January 15, 2024, but the sampling strategy remained unchanged (<u>https://www.socialscienceregistry.org/trials/3509/history</u>).

I4R DP No. 216

fundamental alteration of the study design and its implications for the interpretation of the results is not explicitly discussed in AHI.

4. Classification inconsistencies

4.1 Classification criteria pre-specified vs. published classification criteria

The classification of the villages might also be related to the peculiarity described in the previous section but it deserves a separate examination because of its importance for AHI's main claim. AHI classified 154 villages as government villages and 148 villages as opposition villages, which they then used in their stratification underlying the paper's main result. To trace this classification process, we again compare PAP, DP, and AHI. AHI's default criterion for classification is consistently described across all three documents: official polling station-level election data from the 2001 and 2008 elections. According to AHI, villages are classified as either government or opposition villages if the respective party received more votes in both the 2001 and 2008 elections. This led to a distinct classification "in about two-thirds of the villages" (AHI, p.1310). PAP, DP and AHI are consistent in that the remaining villages must be classified differently. According to AHI, this remaining one-third of villages is classified based on focus group discussions, which their field staff conducted in all villages. PAP and DP mention, somewhat consistently with AHI, that "local knowledge" is used for this classification.

What differs quite substantively between PAP and DP on the one hand and AHI on the other is the description of why the default classification cannot be used for the remaining villages. The PAP and DP explain that those remaining villages cannot be ascribed to one party because they are located in the catchment areas of polling stations with multiple villages, making the attribution of election results to one village impossible. In contrast, AHI report that the remaining villages were categorized differently than the default because their majority vote shifted between the 2001 and 2008 elections (see Table 3).

Table 3: Descri	ption of village	classification	at different sta	ges of the paper

Pre-analysis plan (PAP)	Discussion paper (DP)	Published paper (AHI)
The study area includes 4-5	First, we collected the vote	We selected one village per polling station
constituencies containing	share of the AL and the BNP at	and applied a simple rule whenever
around 800 polling stations	the level of polling stations in	possible to classify villages based on
(PS). Many of these PS consist	the 2001 and 2008 general	previous election results. A village was
of more than one village. []	elections. For polling stations	classified as a government (opposition)
We collect information on the	consisting of multiple villages,	village if the AL (BNP) and its coalition
vote share of the party of the	we complemented this	parties got more votes than the BNP (AL)
current incumbent	information with local	and its coalition parties in the polling
government (AL) and the main	knowledge about the support	station containing this village in both the
opposition party (BNP) at the	for the AL and the BNP across	2001 and the 2008 general elections. We
PS-level in the general	villages within polling stations	could apply this simple rule in about two-
elections in 1996, 2001 and	(talking to village elders,	thirds of the villages in our sample. If the
2008. For PS consisting of	leaders, and party members,	relative electoral strength of the two
multiple villages, we	and verifying from multiple	parties changed from 2001 to 2008, we
complement this information	such sources). We chose 302	used information from focus-aroup-based
with local knowledge about	rural villages from the same	village questionnaires (rather than polling
the support for AL and BNP	number of polling stations.	station-level information) to classify
across villages within PS.	(DP, p.7)	villages. (AHI, p.1310)

It is possible that AHI changed the criterion underlying the classification decision compared to the DP, and that it coincidentally led to similar shares of two third and one third of the villages. But they are likely to be mutually exclusive. The two descriptions would only be consistent if all villages within multi-village polling stations experienced a majority vote change between 2001 and 2008, while all other villages did not. This seems very implausible, and AHI's data also confirm the discrepancy between the two classification approaches. We can demonstrate that the overlap between villages from multi-village polling stations and villages that changed their majority vote is minimal, with only 29 of the 100 villages classified based on focus groups fulfilling both criteria (see Figure A1).

4.2 Inconsistencies between data and maps

It is furthermore puzzling that both the DP and AHI feature maps displaying the sampled villages with identical classification and treatment assignments, despite the different description of the classification process (DP, p.17; AHI, p.1312). We investigate this further by comparing the mapped village classifications and mapped treatment assignments (henceforth: AHI map) with the corresponding data from AHI's reproduction package (henceforth: AHI data). For each of the five sub-districts covered by the study, by visual inspection of the map, we count the number of control villages, policy treatment villages, and legitimacy treatment villages, separately for villages classified as government and opposition, and compare these

numbers to those in the AHI data. This coarse approach misses misalignments *within* the subdistricts and is thus a conservative lower bound.

Table 4 presents the results with deviations in bold. We identify 17 discrepancies across all sub-districts, indicating that the AHI map is inconsistent with the AHI data in terms of village classification, treatment assignment, and total number of sample villages. We cannot rule out that these discrepancies occurred in our visual inspection of the maps, by misassigning villages close to sub-district borders to the wrong sub-district. Yet given the other irregularities in the description of classification process and given the importance for AHI's main claim further examination seems warranted.

Sub-district:	Assa	isuni	Dun	nuria	Ко	yra	Paikg	gacha	Та	ala	То	tal
	map	data	map	data	map	data	map	data	map	data	map	data
Government villages												
Control	16	16	13	14	2	3	12	13	7	6	50	52
Policy treatment	7	8	18	17	5	5	12	12	9	9	51	51
Legitimacy treatment	12	12	15	18	7	5	9	9	7	7	50	51
Opposition villages												
Control	10	11	5	4	14	14	11	10	8	9	48	48
Policy treatment	5	5	17	17	9	9	11	11	8	8	50	50
Legitimacy treatment	12	11	8	8	11	12	13	13	6	6	50	50
Total	62	63	76	78	48	48	68	68	45	45	299	302

Note. Discrepancies in village numbers between the map and the data are shown in bold.

To complement the coarse visual approach, we refine the comparison between AHI data and AHI map by incorporating GPS data from official election data. This allows us to compare treatment status and village classification at the individual village level, revealing further consistency concerns. Figure 2, Panel A shows how we linked treatment status and village classification from the AHI data with polling station locations. Since the AHI data does not include polling station names we obtained official 2018 election results from the *Bangladesh Election Commission*.⁸ Since AHI data also contains some variables from these official election results at the level of individual polling stations, we could use these to successfully match 292 of the 302 villages in the two datasets (the variables are voter turnout, election results for

⁸ As of December 2024, the official election results referenced in AHI are no longer available on the Bangladesh Election Commission website. We used Wayback Machine, a digital archive, to obtain the official election results going back to August 14, 2022:

https://web.archive.org/web/20220814212437/http://www.ecs.gov.bd/page/parliament-election-result

I4R DP No. 216

individual parties, the number of illegal votes, and the electoral constituency). We obtain GPS coordinates for these polling stations from online sources, mostly Google Maps.⁹

While we still cannot locate individual villages from the AHI data on a map, we can locate polling stations. Additionally, we know from the AHI data that 95% of villages actually have a polling station in the village. We combine these pieces of information by overlaying GPS coordinates of polling stations onto the AHI map (Figure 2, Panel B). This final step enables us to directly compare whether the treatment and village classification information (depicted by symbols and colours) in AHI's map aligns with the classification and treatment status we linked to polling stations.

Figure 2: Map reproduction process



In Figure 3 we show an example from *Tala* sub-district. The map overlays survey site information from the AHI map (darker symbols) with polling station information identified by us (lighter symbols). We successfully match polling station information to 44 of the 45 sample villages in the sub-district.

The map confirms the inconsistencies between the AHI map and the AHI data. To exemplify this the figure includes a zoomed-in section of the north-western cluster of villages from the map and a manual matching attempt, linking polling stations to sample villages based on treatment assignment, village classification, and geographic proximity. We find mismatches (marked by red-numbered circles on the map) even when prioritizing matches in treatment

⁹ Polling stations are for the vast majority in schools. Using the school names and the Upazila, we were able to identify almost all polling stations, sometimes relying on additional sources such as Facebook.

assignment and village classification over proximity. All mismatches involve discrepancies in government/opposition classification.

The larger map suggests that the inconsistencies observed in the zoomed-in section extend beyond this part of the sub-district. In fact, we find nine mismatches for the entire sub-district with 45 sample villages and find other non-sampled polling stations that seem closer to sample villages. The map furthermore shows several polling stations that seem either too far away from villages or an implausible number of polling stations can be found near a much smaller number of villages, raising additional questions about the quality of the sampling or mapping. Overall, we find strong indications that either the AHI map, the AHI data, or the description in AHI include incorrect information.



Figure 3: Example of map analysis

5. Conclusion

While we successfully computationally reproduce the results presented in AHI using the replication package, our robustness reproduction identifies a distinct pattern in the main outcome and explores potential explanations. Moreover, our analyses reveal substantial inconsistencies in both AHI's reporting and data. These inconsistencies create ambiguities related to three pivotal elements of the study: village classification, treatment assignment, and sampling. Underlying explanations, we believe, have the potential to alter the theoretical

claims made in AHI and hence the paper's contribution to the literature. In any case, the identified inconsistencies warrant further clarification.

References

Ahmed, F., Hodler, R., & Islam, A. (2020). Voting or abstaining in 'managed' elections? A field experiment in Bangladesh. *CEPR Discussion Paper No.* DP14608

Ahmed, F., Hodler, R., & Islam, A. (2023). Replication package for: "Partisan effects of information campaigns in competitive authoritarian elections: Evidence from Bangladesh" [Data set]. Zenodo. https://doi.org/10.5281/zenodo.10059860.

Ahmed, F., Hodler, R., & Islam, A. (2024). Partisan effects of information campaigns in competitive authoritarian elections: Evidence from Bangladesh. *The Economic Journal*, *134*(660), 1303-1330.

Ahmed, F., Hodler, R., & Islam, A. (2025). Response to Bensch, Rose, Neubauer, Ankel-Peters, and Brodeur.

Association of World Election Bodies. (2025). *Bangladesh*. Retrieved on January 24, 2025, fromhttps://aweb.org/eng/wiki2/kaw05_02_view.do?menuNo=300065&countrySn=4.

Banerjee, A., Duflo, E., Finkelstein, A., Katz, L.F., Olken, B.A. and Sautmann, A. (2020). In praise of moderation: Suggestions for the scope and use of pre-analysis plans for RCTs in economics. National Bureau of Economic Research Discussion Paper No. w26993

Bangladesh Bureau of Statistics; OCHA Regional Office for Asia and the Pacific (2020). Bangladesh administrative level 0-4 boundary shapefiles. Available online at: https://data.humdata.org/dataset/cod-ab-bgd.

Bensch, G., Rose, J., Neubauer, F., Ankel-Peters, J., & Brodeur, A. (2025). A comment on "Partisan effects of information campaigns in competitive authoritarian elections: Evidence from Bangladesh" by Ahmed et al. (2024).

Brodeur, A., Fiala, L., Fitzgerald, J., Kujansuu, E., Valenta, D., Rogeberg, O., & Bensch, G. (2025). A comment on "Improving women's mental health during a pandemic". *IAR Discussion Paper Series* 207.

Election Commission Bangladesh. (2001). Constitution of the People's Republic of Bangladesh: rulesandregulations.RetrievedJanuary24,2025,fromhttps://www.ecs.gov.bd/files/Khdh5RFCFWlp1NOVIdJv3AsAPsvRoFHAOTZBBgZ7.pdf

Fiala, L., Kjelsrud, A., Kujansuu, E., & Brodeur, A. (2025). Comment on "Food insecurity and mental health of women during COVID-19: Evidence from a developing country" by Rahman et al. Available at https://osf.io/sgyaj.

Giné, X., & Mansuri, G. (2018). Together we will: Experimental evidence on female voting behavior in Pakistan. *American Economic Journal: Applied Economics*, *10*(1), 207-235.

Hodler, R. & Islam, A. (2018, 2024). Voting or abstaining in "managed" elections? An RCT in Bangladesh. AEA RCT Registry. December 20 (updated on January 15, 2024). https://doi.org/10.1257/rct.3509-1.0

Kilic, T. & Moylan, H. (2016). Methodological experiment on measuring asset ownership from a gender perspective. World Bank, Washington, DC, URL: <u>http://hdl.handle.net/10986/33653</u>.

Kjelsrud, A., Kotsadam, A., Rogeberg, O., & Brodeur, A. (2025). A comment on "Raising health awareness in rural communities: A randomized experiment in Bangladesh and India" by Siddique et al. (2024). *IAR Discussion Paper Series* 208.

Masselus, L., & Fiala, N. (2024). Whom to ask? Testing respondent effects in household surveys. *Journal of Development Economics*, 168, 103265.

Appendix

Appendix A: Complementary results



Note: The figure shows the overlap between FGD villages based on the descriptions in AHI and PAP/DP. The AHI data includes the 2001 and 2008 election data linked to each sample village. Using this data, we searched for exact matches in the 2001 or 2008 election results – examining total number of voters, government party votes, and opposition party votes—always within the same electoral constituency. Cases where multiple villages share identical election results are strong indicators that they belong to the same polling station. This expectation is reinforced by the fact that no such exact matches appear in the 2018 election results that we retrieved for all 595 polling stations in the four constituencies studied (see footnote 8), even when considering only total votes and votes for the governing party. Note also that we find a high number of villages likely linked to identical polling stations, because this refers to the 2001 and 2008 polling stations used for the classification, when much fewer polling stations existed than in 2018 (Bangladesh Election Commission 2001; Association of World Election Bodies 2025).